DOCUMENT RESUME

ED 144 983 TM 006 534

TITLE	Violations of Human and Civil Rights: Tests and Use of Tests. Report of the Tenth National Conference on Civil and Human Rights in Education, February 18-20, 1972, Washington, D.C.
INSTITUTION	National Education Association, Washington, D.C.
SPONS AGENCY	National Education Association, Washington, D.C. Center for Human Relations.
PUB DATE	Feb 72
NOTE	35p.
AVAILABLE FROM	National Education Association Order Dept., Academic Building, Westhaven, Connecticut 06516 (Stock Number 0556-2-06, $1.00); cassette tapes of condensations of conference speeches available from same address (Stock Number 0550-3-06, $28.50)
EDRS PRICE	MF-$0.83 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS	Bilingual Students; *Civil Rights; Cultural Differences; *Educational Opportunities; Elementary Secondary Education; *Minority Group Children; Public Schools; Self Concept; *Standardized Tests; *Student Rights; *Test Bias; *Testing Problems; Test Interpretation

ABSTRACT

Major speeches presented at the conference on human and civil rights in education are summarized. Issues discussed at the conference include: the extent to which standardized testing infringes on the rights of students and educators; the existing uses of educational testing in public schools; and ways to eliminate misuses of tests by educators, researchers, and the testing industry. Summaries of small group discussions are represented on test bias, use of tests for educational administration and employment counseling and the impact of inaccurate test data on student self concept. Recommendations of these discussion groups are summarized. Results of a questionnaire on testing are also included. (GDC)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

OF HUMAN AND CIVIL RIGHTS:

TESTS AND USE OF TESTS

Report of the Tenth National Conference on Civil and Human Rights in Education

TM006 534

violations OF HUMAN AND CIVIL RIGHTS: TESTS AND USE OF TESTS

Report of the Tenth National Conference on
Civil and Human Rights in Education

February 18-20, 1972
Washington, D.C.

Sponsored by the NEA Center for Human Relations

National Education Association
Washington, D.C. 20036

Cassette tapes of the conference speeches (condensations) are available from NEA Publishing, 1201 Sixteenth Street, N.W., Washington, D.C. 20036

Copyright © 1972
Library of Congress Catalog Card No. 72-86306
NEA Stock No. 381-12004

Contents

	PAGE
Introduction	4
Conference Schedule	6
Sociocultural Factors in the Education of Black and Chicano Children	
Keynote speech by Dr. Jane R. Mercer	8
Bias in Testing	
Speech by Dr. William F. Brazziel	11
Use of Tests: Educational Administration	
Speech by Dr. José A. Cardenas	13
Use of Tests: Employment and Counseling	
Speech by Dr. Thelma Spencer	15
Misuse of Tests: Self-Concept	
Speech of Dr. Robert L. Williams	17
The S.H.A.F.T. Test	20
Wrap-Up	
Speech by Dr. Dwight Allen	23
Recommendations	26
Testing Bilingual Children	
Statement by Bay Area Bilingual Education League	28
Appendix	
Testing Opinionnaire	29

Introduction

The tenth national NEA-CHR conference on civil and human rights in education was held in Washington, D.C., February 18-20, 1972. This informational conference on testing constituted an important part of NEA's ongoing program to protect the human and civil rights of educators and children and to seek quality education for all Americans.

The conference provided a national forum for airing issues relating to the use of standardized tests and their effects upon educational opportunities. It focused on three major questions:

1. To what extent do standardized tests and their use infringe on the civil and human rights of educators and students?
2. What is and what should be the use of standardized tests in public education?
3. What actions are needed to eliminate misuse of tests by researchers, psychologists, school administrators, counselors, teachers, and the testing industry?

Testing and the use of standardized tests have long been considered an essential part of the evaluation process in schools. Tests have provided tools for individual students and teachers to compare and contrast performance levels. Tests have the potential for assisting in individualized instruction programs, and they provide a crude measure for administrators, teachers, students, and parents to judge the effectiveness of the educational experience. However, it has become increasingly clear that tests and the use of tests have been dysfunctional and harmful for many children.

Conference participants assembled in four forum groups to explore four ways that tests limit children's opportunities:

A. By measuring children against criteria inappropriate for their cultural and educational needs
B. By being used as a method for "sorting out" children into tracks or groups that maintain limited horizons and reinforce negative self-concepts
C. By restricting opportunities for higher education and employment of individuals who might otherwise be successful in their chosen endeavors
D. By contributing to the "self-fulfilling prophecy" which ensures low-level achievement for children who score low on tests.

Designed to find solutions to these issues and to examine the benefits and liabilities of standardized tests, the tenth national CHR conference gave psychologists, counselors, teachers, administrators, students, and parents the opportunity to obtain information, to engage in dialogue, and to determine the implications of the problem as it relates to their professional roles.

This report contains summaries of the major speeches given during the conference and attempts to reflect the conference mood and the valuable contributions made by participants. The four forum groups—Bias in Testing, Use of Tests: Educational Administration, Use of Tests: Employment and Counseling, and Misuse of Tests: Self-Concept—were assembled to ensure fuller coverage of all the issues related to the use of tests. Small discussion groups were charged with the responsibility of discussing specific questions and coming up with recommendations. The concern of participants was evident in their commitment to this task. Many small groups continued their meetings into the wee hours of the morning to complete their recommendations.

Each small group presented its recommendations to its forum, and the forum recommendations were then given to the total conference. These recommendations are summarized on pages 26-28. They do not represent official NEA positions.

It is our hope that the conference and this report will stimulate the growing dialogue on the use of tests and lead to reform which will contribute to the achievement of truly equal opportunity for quality education.

George W. Jones, *Director*

Conference Schedule

Registration

February 18
1:00 p.m.-10:30 p.m. — *Upper Lobby (Federal Wall)*

February 19
8:00 a.m.-2:00 p.m. — *Upper Lobby*

February 20
9:00 a.m.-12 noon — *Upper Lobby*

Display Area

February 20 only — *Capitol Terrace*

NEA Office and Press Room — *California Room*

Friday, February 18, 1972

4:00 p.m.-5:30 p.m.

Orientation for Discussion Leaders — *New York Room*

8:30 p.m.-10:30 p.m.

Opening Session — *Congressional and Senate Rooms*

"Sociocultural Factors in the Education of Black and Chicano Children"

Presiding
Wade Wilson, Chairman
NEA Human Relations Council
Pres., Cheyney State College
Cheyney, Pennsylvania

Greetings
Samuel B. Ethridge, Asst. Exec. Secy.
NEA Teacher Rights Division
Washington, D.C.

Introduction of Speaker
Lois Wilson, Asst. Exec. Secy.
New York State Teachers Association
Albany, New York
and First Vice Chairman
NEA Human Relations Council

Keynote Speaker
Jane Mercer
Associate Professor and Research Specialist
University of California
Riverside, California

10:30 p.m.-11:30 p.m.

Reception — *Federal, Foyer II, and South American Rooms*

Saturday, February 19, 1972

9:00 a.m.-10:30 a.m.

FORUMS

"Uses and Misuses of Tests"

A. Bias in Testing
Congressional Room

Presenter
William F. Brazziel
Professor of Higher Education
University of Conn.
Storrs, Conn.
Reaction by panel

B. Use of Tests: Educational Administration
Senate Room

Presenter
José Cardenas
Superintendent
Edgewood I.S.D.
San Antonio, Texas
Reaction by panel

C. Use of Tests: Employment and Counseling
South American Room

Presenter
Thelma Spencer
Assistant Director
National Teacher Examinations Program
Educational Testing Service
Princeton, New Jersey
Reaction by panel

D. Misuse of Tests:
Self-Concept

Federal Room

Presenter
Robert L. Williams
Professor of Psychology
and Director of Black Studies
Washington University
St. Louis, Missouri
Reaction by panel

10:45 a.m.-12:15 p.m.

WORK GROUP MEETINGS

Rooms to be announced at Forum Sessions

Discussion and Formulation of Recommendations

12:30 p.m.-2:00 p.m.

LUNCHEON SESSION

Presidential Room

Presiding
Al Colum
Second Vice Chairman
NEA Human Relations Council
San Diego, California

Greetings
Donald E. Morrison
NEA President
San Diego, California

Sam M. Lambert
NEA Executive Secretary
Washington, D.C.

2:30 p.m.-4:30 p.m.

WORK GROUP MEETINGS

Continued in same work group rooms

4:30 p.m.-7:30 p.m.

FREE TIME

Rooms available for special interest groups

DINNER—OPEN

7:30 p.m.-8:30 p.m.

WORK GROUP MEETINGS

Continued in same work group rooms

Sunday, February 20, 1972

9:00 a.m.-10:30 a.m.

FORUMS
Presentation of work group recommendations

A. Bias in Testing

Congressional Room

B. Use of Tests: Educational Administration

Senate Room

C. Use of Tests: Employment and Counseling

South American Room

D. Misuse of Tests:
Self-Concept

Federal Room

10:45 a.m.-12 noon

GENERAL SESSION

Presiding — *Congressional and Senate Rooms*
Margaret Fischer
Forum Reports

12:30 p.m.-2:00 p.m.

LUNCHEON AND CLOSING SESSION — *Presidential Room*

Presiding
Wade Wilson

Wrap-Up
Dwight Allen
Dean, School of Education
University of Massachusetts
Amherst, Massachusetts

Keynote speech by Dr. Jane R. Mercer, associate professor of sociology, University of California, Riverside, California.

Studies dating back to the 1930's have demonstrated the cultural biases inherent in IQ tests and other standardized achievement measures. Yet clinicians have continued to interpret children's performances on these tests as if there were no biases and have never systematically taken sociocultural differences into account when interpreting the meaning of a particular child's score. Consequently, we find many children in classes for the mentally retarded whose adaptive behavior in non-academic settings clearly demonstrates that their problems are school specific and that they are not comprehensively incompetent.

Disproportionately large numbers of Black, Chicano, and probably Puerto Rican children are labeled mentally retarded by the public schools. In California, the rates for placing Chicano and Black children in classes for the mentally retarded are two to four times higher per thousand than for English-speaking Caucasian children, or Anglos.

The public schools label as retarded a large number of children who are not so regarded by their families, neighborhoods, churches, or other community organizations. We asked

241 organizations in a Southern California city for information on each retarded person they were serving. The public schools listed far more retardates than any other formal organization, shared their labels with more other organizations, and labeled as retarded more persons with IQ's above 70 and with no physical disabilities. There were 4½ times as many Chicanos and twice as many Blacks in public school classes for the mentally retarded as would be expected from their proportion in the population, and only half as many Anglos as would be expected. The Black and Chicano children in these classes had higher IQ's and fewer physical disabilities than the Anglo children. While we found no evidence that these ethnic disproportions resulted from a conscious policy of discrimination, the labeling process is clearly Anglocentric.

We then sought to identify the aspects of the clinical assessment process that produce ethnic disproportions, by testing a representative sample of 6,907 persons in the community. We used the American Association for Mental Deficiency's definition of mental retardate as a person subaverage in both general intellectual functioning and adaptive behavior and developed a series of 28 age-graded scales to measure adaptive behavior. We also used standardized measures of intelligence, mainly the Stanford-Binet LM and the Kuhlmann-Binet.

We found that the educational institutions' definition of mental retardates as those with IQ's of 79 or below—the lowest 9 percent of the population—is one factor producing ethnic disproportions in the labeling process. We concluded that a 3 percent cutoff—IQ below 70—is most likely to identify persons in need of special assistance and least likely to stigmatize those who perform a normal complement of social roles.

Most psychologists give only an IQ test when making assessments of mental retardation. However, we found that 60 percent of the Chicanos and 91 percent of the Blacks in our sample who had IQ test scores below 70 passed the adaptive behavior measure, while none of the Anglos with IQ's this low were performing normally in their social roles. The IQ test is obviously not a valid predictor of social role performance for Chicanos and Blacks, although it seems to do a good job for Anglos. Schools should adhere to the AAMD definition of mental retardation and develop a systematic method for measuring adaptive behavior as well as IQ in making psychological assessments. A child should have to fail both criteria before being labeled mentally retarded. When we followed this procedure, ethnic disproportions were reduced but still not completely eliminated.

The IQ tests now being used by psychologists are, to a large extent, Anglocentric. We found that about 32 percent of the differences in IQ test scores in a sample of 1,500 Black, Chicano, and Anglo elementary school children in California could be accounted for by differences in the sociocultural characteristics of their families. Unfortunately, most psychologists treat a score as a score. When social background was held constant, there was no difference between the measured intelligence of Mexican-American and Black children and the Anglo children on whom the test was standardized. We concluded that diagnostic procedures in the public schools must be broadened to reflect the pluralistic nature of American society and must involve securing information beyond that ordinarily used in public school assessment.

A major concern of parents we interviewed in our studies was the stigmatization of their children. Their children were ashamed to be seen entering the mentally retarded room and dreaded receiving mail that might bear compromising identification. The parents were also concerned about the quality of the educational program in the self-contained special

education class. Their children were not taught to read as they would be taught in the regular classes, and many saw the program as a "sentence of death." We followed a group of 108 children in special classes for several years; only one in five ever returned to regular classes. The others aged out, dropped out, or were sent to other special programs.

It would be a tragedy, however, if special education programs were jeopardized because of inadequacies in assessment procedures and programing. I believe that there are viable alternatives to present practices without resorting to dumping special education students and cutting special education funds. School psychologists should be required to enlarge the scope of information they use in making educational decisions by regularly and systematically studying students' adaptive behavior in nonschool situations. If a child performs adequately in these settings, his problems are school specific, and he needs special tutoring, programed learning, cross-age teaching, and remedial reading, rather than a self-contained classroom. School psychologists should also secure information about a child's sociocultural background to use in interpreting his IQ test score and developing pluralistic norms. The child's performance should be compared not only with the performance of the general population, which is composed primarily of Anglo children, but also with the performance of other children from his own sociocultural background.

I do not agree with those who say we mus stop all educational labeling or IQ testing. Ou problem has been that our labels are too fe and too crude. We need a more sensitive sys tem for identifying children in need of speci education and a continuum of special educa tion programs carefully targeted for childre with specific needs. One of the most distress ing developments in some regions has bee the precipitous reassignment of many childre to the regular classroom from self-containe classrooms with no continuation of speci services. These children must continue t receive special education, and that mone must continue to be provided. Financial sup port and the effort of special education teach ers should be redirected toward providing wider variety of special services to keep chi dren in the educational mainstream and t educate each to his maximum potential.

This is a condensation of Dr. Mercer's speec The full text is available from the NEA Cent for Human Relations; single copies free whi they last.

A more comprehensive and detailed analys and report of the referral process, the clinic testing process, the assumptions and inferenc of the clinician, and the pluralistic evaluatio process discussed herein appear in Dr. Mercer forthcoming volume entitled Labeling the Men tally Retarded. *Berkeley: University of Californ Press, 1972.*

Forum A — Bias in Testing

Speaker: Dr. William F. Brazziel, professor of higher education, University of Connecticut, Storrs, Connecticut.

Thousands of minority children are denied equal access to quality education each year because of flaws in the testing apparatus of our schools. This situation, which is both illegal and immoral, is becoming less and less tenable. At least 20 class action suits around the country are seeking to force school districts to cease and desist in the inaccurate testing of minority children. In the biggest suit, in California, the NAACP and several civic groups are seeking a dissolution of classes for the retarded and a moratorium on testing until better, more precise instruments are devised.

The minority child's testing problems begin on his first day at school, when he has to take an imprecise test and have his score recorded on his cumulative record. Should it be less than 100, his teachers will not work diligently with him and will give him less than his share of attention and assistance. He will get more than his share of slights and indignities as he moves through the schools and will be denied access to a college preparatory curriculum.

More and more people are losing confidence in the schools, and the spectacle of a testing apparatus in disarray will do little to restore this confidence. The $300 million testing industry must come up with better instruments. The schools must eliminate injurious instruments, and psychometrists and teachers must be retrained to ensure that tests become a

part of the solution rather than the problem in American schools. Resistance to this reform movement will come from long-time test consultants, conservative school people, racists, teachers of teachers, and the psychometric profession.

Henry Dyer, vice-president of the Educational Testing Service and the dean of American psychometrists, says that IQ tests are the most useless source of educational controversy ever invented and that schools which have not yet dropped them should do so forthwith. He notes that a more sophisticated testing apparatus could be developed for schools with a heterogeneous population, but that the continuation of IQ tests would preclude this. Dyer is right. The very best test in America was standardized on 4,400 children, most from California suburbs. Neither minority children nor children from the Southeast were included in the sample.

Dyer suggests more school programs based on the philosophy of Jean Piaget. In the Piagetian school, teachers, tests, and curriculum are viewed as resources to maximize each child's development. Instead of being slapped in the face with a biased IQ test, a child takes a battery of sophisticated tests designed to ascertain his language development, comprehension, symbol manipulation, discriminant analysis, and other skills. None of his scores are recorded in his cumulative folder. Criterion referenced tests are used instead of the norm-based achievement tests, so child racing is eliminated. In the Piagetian school, parent conferences and skill sheets replace report cards, and continuous progress learning replaces promotion from grade to grade.

Piagetian teachers sometimes give environmental or culture-specific tests, in which familiar concepts from the child's neighborhood are used to ascertain his ability to think. A Mississippi child, for example, might be asked to match singletree, lespedeza, sweetmilk, tedder, dasher, hamestrap, blue tick, and walker, instead of sonata and bas-relief. The content of any test is irrelevant in the Piagetian school.

Culture-specific tests are not new, but the testing corporations have found it unproductive to develop tests for each of the 40 or so cultural groups in this country. The government and the corporations should have such tests available in a few years, but in the meantime school systems and individual teachers will have to make their own. There is no great mystery about test making; over 10,000 tests are on the market. Like the struggle to get publishers to market integrated textbooks, this movement will probably have to resort to teacher and school system efforts to prime the pump.

The situation concerning criterion referenced tests is better. These achievement tests, which measure only what has been taught—in a particular module, by a particular teacher, and in a particular time span—come with such model programs as Distar and individually prescribed instruction. Their value lies in their elimination of the need to have losers in the testing game who make it possible for others to succeed. They focus on growth and behaviorally oriented goals and so benefit both low-income and middle class children.

Nothing that is wrong with the testing programs in American schools is immune to hard work, imagination, and nerve. The denial of equal access to quality education is criminal, and the public will not tolerate it. If a few in our midst are resistant to change, we shall find ourselves caught up in an embarrassing maze of court suits, reduced budgets, performance contracts, school vouchers, and steadily eroding public confidence. Our job is to make sure that this does not happen and that American schools are modified to serve well all the children of all the people.

This is a condensation of Dr. Brazziel's speech. The full text is available from the NEA Center for Human Relations; single copies free while they last.

Forum B—Use of Tests: Educational Administration

Speaker: Dr. José A. Cardenas, superintendent, Edgewood Independent School District, San Antonio, Texas.

The purpose of evaluation and testing is decision making. Tests—of intelligence, achievement, and personality—provide information for making decisions about continuation, promotion, graduation, special assignment and placement, diagnosis and prescription, student feedback, motivation, and evaluation. I will focus on the intelligence test.

The Assumptions

The most serious problem in the assessment of intelligence and the use of intelligence tests is the assumptions that are made. One assumption is that intelligence is intangible and not directly measurable. Another is that intelligence can be indirectly measured by assessing some form of behavior or performance that is assumed to be solely dependent on intelligence and not related to other variables such as related understandings, prior learnings, and motivation. A third assumption is that the evaluation of behavior and performance requires written, oral, or nonverbal interaction between tester and testee. A fourth is that test results give little or no information unless a comparison is made between the performances of the testee and the norm group. We also assume that the charactistics of the two are compatible.

Research indicates, however, that more is unknown than is known about intelligence and that the assumptions and methods of testing are not always valid. Performance can be based on more than the one variable of intelligence. Our assumption that all other factors are equal or are nondependent variables that have nothing to do with intelligence is false. Language facility, reading ability, and cultural compatibility all influence test scores.

Too many intelligence tests assume that all children have had common experiences—for example, that they are all familiar with snow. We haven't had a snow holiday for the last five years in the Edgewood school district, and most of the kids have never seen snow, yet some group intelligence tests ask about the use of a sled. One question on the WISC asks, "If your mother sends you to the store for a loaf of bread and there is none, what do you do?" The child who answers, "I go back home," is considered to be intellectually inferior to the child who says, "I go to another store." However, in rural areas, there is no place else to go. In some families, the parent and not the child is supposed to make such a decision. During my youth, I was sent for tortillas, and the purchase of a loaf of bread was unheard of in my house until I was 15 years old.

Testing methods can also be incompatible with a culture. For example, some tests em-

phasize competitiveness, but Mexican American children perform better in a cooperative situation.

Our assumption that an individual has the ability to verbalize can harm children with physical disabilities. A child who stammers will probably score low on a verbal or language test even though his disability is unrelated to his intelligence. Motivation is assumed not to be a factor influencing a child's performance on intelligence tests, yet a test may be highly motivating for some children and totally inhibiting for others. We also assume that the testee-tester relationship is not a dependent variable, yet many studies indicate that Black and Mexican American pupils perform better on intelligence tests when they are administered by Black and Mexican American administrators. Score reliability is, likewise, assumed, but how smart you are is more dependent upon who scores your test than upon your intelligence or your performance.

The Dysfunctional Responses

The second major problem in testing is the dysfunctional responses we sometimes use to try to remedy our invalid assumptions. It is simplistic to give a Puerto Rican child a Spanish translation of the WISC or the Peabody Picture Vocabulary Inventory. Some English stimulus words become Spanish paragraphs when translated properly. I have never heard a one-word translation of "cream puff," yet this word is on the Spanish version. Also, the Spanish equivalent of an English word may be on an entirely different level of difficulty. No matter how good the Spanish translation of a test, it must also take into account the regional variation of the language. When I once told a test translator that her translation must be regionalized to be valid, she protested the expense and said, "If the kids don't know this type of Spanish, that's their problem." So Spanish-language tests can be just as invalid as English-language tests for Spanish-speaking children.

Bilingual children who take Spanish-language tests receive no credit for their knowledge of English and, in fact, are penalized. The failure of a bilingual child to respond to the stimulus word *mariposa* supposedly indicates lack of intelligence, even though he may be familiar with the word and concept of *butterfly*. When bilingual children are administered a Spanish-language Peabody after having taken the English version, they select the same responses they made in English, even though they now see that some of them are wrong. Even using bilingual tests does not solve the problem; scores have been shown to vary according to whether the child's dominant language appears first or second on the page.

Recommendations

We must protect all children against invalid testing. We must reeducate educational personnel and perhaps discontinue intelligence testing, at least until the reeducation is complete. We must develop new and functional techniques for measuring intelligence and establish different criteria for making decisions. Above all, measures should be taken by the National Education Association and organizations of school administrators and counselors to protect children from the invasion of privacy through testing.

This is a condensation of Dr. Cardenas' speech. The full text is available from the NEA Center for Human Relations; single copies free while they last.

Forum C—Use of Tests: Employment and Counseling

Speaker: Dr. Thelma Spencer, director, Teacher Education Examination Program, Educational Testing Service, Princeton, New Jersey.

Tests are widely used in counseling, but their validity depends on the interpretation of scores and on the situational appropriateness of the tests. The interpretation of scores often indicates the poor quality of the counseling to which students are exposed.

> . . . [A] young girl had received honors in junior high school and top scores in every standardized test she took. As she was about to graduate she was given an interest inventory. According to her responses, the girl's major interest was in clerical work.
>
> The school guidance counselor met with the girl and informed her she should become a secretary. Not only did the counselor tell her she would be happier in a commercial course in high school, he insisted that she would be unable to cope with the intellectual demands of a liberal arts college. In his ignorance the guidance counselor had confused a questionnaire that supposedly shows what people *like* to do with aptitude tests that attempt to measure what people *can* do.[1]

Bad counseling can take other forms, as in this incident related by a New York high school student:

> . . . [T]he people in my section were never told about the PSATs. . . . The kids in the Honor sections took those PSATs as a matter of course. In fact, the teachers strongly urged that they take those PSATs in order to get some practice for the next year's SATs. And I didn't take a PSAT until I got into my senior year, and then I found out I had to take an SAT.

[1] Black, Hillel. *They Shall Not Pass*. New York: William Morrow & Company, 1963. p. 167.

> That sticks in my craw. We were all students and we supposedly were all aspiring to college.[2]

In both of these cases the system worked against the student, but the first student is representative of the type for whom the system is supposed to work. The second is representative of those for whom the system works very seldom.

Tests also mediate against students. Some children are tested more than they are taught, and curriculum is often based on what can be measured by existing instruments. Too many counselors equate a test score with the intangibles of personality—motivation, desire, and ambition—and then ignore personality in favor of a preconceived notion about the individual being counseled. "The best of our tests," as Oscar Buros says, "are still highly fallible instruments which are extremely difficult to interpret with assurance in individual cases."[3] Many counselors and teachers victimize youngsters by equating the individual with the norm and the testee group with the norm group.

Some people say that the evolution of a stable examination system has helped create our much vaunted stable educational system, but while the SAT may permit admissions officers to select students most similar to the norm group, those students come from schools whose programs are based on what the colleges offer. The implications of this vicious circle are staggering: too many poor and minority students who fail to meet admissions criteria are guided into general, vocational, and commercial classes.

Sometimes a student will succeed despite a counselor's or teacher's doubts about his ability to do so. Many students never even see a counselor, or if they do, the counselor is a disciplinarian and attendance taker, not a helper and adviser. For the majority of students, adequate counseling and guidance are myths.

Using test scores to determine employability is another gross misuse of tests. For example, the National Teacher Examinations have been used by some to determine who should be retained when a southern school district is under court order to dissolve its dual system. When the school system in Columbus, Mississippi, required a cutoff score of 1,000 on the NTE for retention in the unified district, eight Black teachers were not rehired. In Starkville, Mississippi, the school district used the Graduate Record Examination to determine teacher retention, although it granted provisional status to teachers with NTE scores of 500 on both the common and teaching field sections. These districts equated a test score with competence in the classroom. In April 1971 Judge Orma Smith ruled that personnel selection based on NTE and GRE results was discriminatory, as more whites than Blacks scored above the cutoff point in both districts. In *Griggs et al. v. Duke Power Company* the U.S. Supreme Court ruled that tests for employability violate the 1964 Civil Rights Act "when the rate of rejection is higher for Negroes than for whites and there is no showing that the passing of such tests is significantly related to the successful performance of the job."

Test scores are guides only, and the NTE score is merely another piece—by no means the most important piece—of information about a person. This test, or any test, is only as good as the people who use it.

This is a condensation of Dr. Spencer's speech. The full text is available from the NEA Center for Human Relations; single copies free while they last.

[2] Wasserman, Miriam. *School Fix, NYC, U.S.A.* New York: Outerbridge and Dienstfrey, 1970. pp. 155-56.
[3] Quoted in Black, *op. cit.*, p. 260.

Forum D – Misuse of Tests: Self-Concept

Speaker: Dr. Robert L. Williams, director, Black Studies Program, and professor of psychology, Washington University, St. Louis, Missouri.

The problem of testing Black children is very serious. Biased tests are not just a violation of civil rights, but are a form of Black intellectual genocide. The whole American educational system is unfair, and the argument against tests is used as one instrument to open the door to change the whole system.

What do we mean by intelligence? Is it what the intelligence tests measure? Is it a global capacity to deal with one's environment? I offer the rubber band theory to illustrate my definition of intelligence. A rubber band will stay in its relaxed state unless stretched to its capacity by an outside force. Genetics or heredity determines an individual's potential stretch, and the environment determines the extent to which he reaches this potential.

Test items drawn from white culture penalize Black children, whereas items drawn from Black culture penalize whites. I have developed the BITCH test – Black Intelligence Test of Cultural Homogeneity – with items drawn directly from Black culture (to be published by May 1, 1972). A child who knows Malcolm X's birthday or the date of his assassination shows as much intelligence as the one who knows Washington's birthday. I've never seen the word *pick* illustrated on the WISC – only *comb*, which is something I can't use.

The three criteria for a test—validity, reliability, and standardization—exclude Black people. A test is valid when it measures what it intends to measure. Currently used ability tests do not measure Black intelligence. If a test asks, for example, "What should you do if you find a purse with five dollars in it?" a Black child will say, "Keep it"—a culturally determined response. He will say what his environment has dictated that he say, but on the standardized test, he will be marked zero.

Reliability means test consistency—i.e., a test will yield the same score, or rank an individual in the same place, each time. Since the standardized tests are scored subjectively, and since they validate only mainstream cultural responses, they cannot be reliable.

Standardization refers to the extent to which the sample on which the test is based represents the people who will take it. Several of the major ability tests excluded Blacks, Mexican Americans, and Puerto Ricans from their standardization samples. The Stanford-Binet, WISC, and Peabody systematically excluded Blacks from their samples. If a test is not standardized on a particular group, it probably does not represent that particular group and should not be used on its members. Standardization is one reason for the 15-point difference in IQ between Black and white kids. The discrepancy means simply that the test is biased.

Arthur Jensen's research has repeatedly shown that tests are biased—that Black and Chicano kids who have IQ's in the 60-to-80 range score much higher on his Learning Test than white kids with the same IQ. Jensen's interpretation is that Blacks show more associative or Level I learning than whites; if you ask Black and white kids to recite six digits backwards and forwards, Black kids do better than whites with the same IQ. An alternative interpretation is that the biased IQ underestimates the Black child's ability and indicates that he is clearly superior to the white kids who scored within the same range on the test,

because the test was standardized on the whites.

Some people argue, "The tests do exactly what they are supposed to do. They predict scholastic success." Ability tests (X) predict a criterion (Y), such as a child's performance in the classroom. The hidden fallacy is variable Z, which might be unfairness, motivation, anxiety – anything that influences X (test scores). A fair test and a fair criterion will produce a high correlation between X and Y: white people who do well on tests do well in school. The unfair WISC and a fair classroom will produce a low correlation between X and Y: a Black child who does poorly on the test will do well in the classroom. Another combination is a fair predictor, such as the Davis-Eels Games test, and an unfair criterion – the culturally biased classroom. This combination will also yield a low correlation: Black kids will do well on the test and poorly in the classroom. With an unfair predictor and an unfair criterion – the classic situation for the Black child – the correlation is high: the Black child who does poorly on the test also does poorly in the classroom.

After I administered the WISC test to about 500 Black kids, I then gave the BITCH test. Of the 420 children in the low WISC group, 75 to 80 percent scored high on the BITCH. I still have to examine other criteria to see how well the BITCH scores correlate with scholastic performance, but at least I know that most of the Black children who scored low on the WISC are neither educationally mentally retarded nor in the borderline defective range.

At least four court suits are now pending on the use of standardized tests in San Francisco, San Diego, Boston, and St. Louis because they violate a child's constitutional rights under the Fourteenth Amendment.

The Boehm's test of 50 basic concepts is clearly written for white folks. It asks the child to select the picture that shows "behind the couch" or "under the table." A Black child does not say "behind" but "in back of" – not "under" but "up under." We are now rewriting the instructions to that test to see if children understand the concepts in Black English. Black and white children can have the same cognition but communicate it differently. To the cognition "few," a Black child might say, "Well, that's not a whole bunch of them." Only the vocabulary is different, and a difference is not a deficiency. You don't evaluate Black people in terms of how white they are, but this is what the tests have done. They do not measure Black ability.

Eliminating and inhibiting intelligence early in life is the best way to keep Blacks out of the system. I would opt for talking to children in the dialect they understand. If a child can understand what you are asking of him on a test, he will probably master the task. You cannot expect an individual who has not been exposed to German or French to understand these languages. This does not mean that he lacks the capacity to learn German or French, but that he lacks that particular exposure.

Black parents should be concerned about both the predictive variable – the test – and the criterion variable – the classroom. I think whites should also be vitally concerned. Brother Charlie Mingus said, "When they came and took the Catholics, I did not complain, because I am not a Catholic. When they came and took the Unionists, I did not complain, because I am not a Unionist. When they came and took the Panthers, I did not complain, because I am not a Panther, but then one day they came and took me." I don't think we should let another generation pass in this country that knows all about extravehicular space activity, atomic physics, and all of these highfalutin things, but does not know what a human being is.

This is a condensation of Dr. Williams' speech. The full text is available from the NEA Center for Human Relations; single copies free while they last.

The S.H.A.F.T.* Test

A highlight of the conference was the participation of more than 40 high school students, who reacted to the speakers, joined in the discussion groups, and wrote and administered the S.H.A.F.T. test.

The students from Callanan Junior High School in Des Moines, Iowa, told conference participants to keep their tests face down until told to start. They would be given 10 minutes to answer 22 multiple choice questions about student culture that were standardized on ninth grade students from Des Moines. Student proctors enforced the no-talking, no-peeking rules, although from time to time the educators asked for erasers, wanted to know the time, and called out, "Hey, Teach! Where's the pencil sharpener?"

The test scores formed a perfect bell-shaped curve—a handful of the 650 conferees scored high, most were average, and a few blew it. Besides the public exposure—conferees had to wear black, red, or yellow armbands depending on their scores, or go bare armed for scoring too low—participants were told that their scores would be made part of their cumulative folders to haunt them for life.

*S.H.A.F.T. stands for "Student's Hype Arranged for Teachers."

S.H.A.F.T.

(Student's Hype Arranged for Teachers)

1. What is the slang word used to describe a blemish?
   a. Zilch
   b. Arg
   c. Zit
2. The author of the book *Right On* is
   a. Julian Bond
   b. Jerry Rubin
   c. Iman Baraka (LeRoi Jones)
3. What are waffle stompers?
   a. Pancake chef
   b. Snowshoes
   c. Ice cream sandwiches
4. *Steal This Book* was written by
   a. Allen Ginsberg
   b. Eldridge Cleaver
   c. Abbie Hoffman
5. Who wrote the song "Purple Haze"?
   a. Jimi Hendrix
   b. Aretha Franklin
   c. James Brown
6. For what purpose would you use a roach clip?
   a. To keep ladies' blouses closed
   b. To hold the end of a reefer
   c. To get rid of bugs
7. Who wrote *Love Story*?
   a. Ryan O'Neal
   b. Henry Mancini
   c. Erich Segal
8. The author of *Alice's Restaurant* is
   a. Bob Dylan
   b. Alice Cooper
   c. Arlo Guthrie
9. What can you get at "Alice's restaurant"?
   a. Soul food

b. Storybooks for children
c. Everything you want

10. "Ripple" is
    a. Rumor in a faculty
    b. Cheap wine
    c. A game of chance

11. The term *rip off* means
    a. To tear
    b. To steal
    c. To cop out

12. What rock group sang the anti-drug song, "The Pusher"?
    a. The Who
    b. Steppenwolf
    c. Blood Rock

13. "Tommy" is
    a. A British cop
    b. A fast sports car
    c. A rock opera

14. "Make tracks" means
    a. To inject dope
    b. To burn rubber
    c. To split

15. To "crash" is to
    a. Have an accident
    b. Come down from the use of drugs
    c. Lose all your money

16. Who were the originators of *Jesus Christ Superstar*?
    a. Kado & Ragni
    b. Rice & Webber
    c. Brewer & Shipley

17. *Lenny* is a play on the life of
    a. Lenny Bruce
    b. Lenny Bernstein
    c. Lenny Brezhnev

18. "Getting off" means
    a. To feel the effect from the drugs you have taken
    b. A vacation from work
    c. To cease the addiction of heroin

19. A "hit" is
    a. A robbery
    b. An internal dose of drugs
    c. A very popular teenager

20. Hash is
    a. Cheap opium
    b. A mixture of pep pills
    c. A resin from marijuana

21. A "hemmie" is
    a. A souped-up engine
    b. An Ernest Hemingway short story
    c. A shirt that has been shortened

22. An ounce of marijuana is referred to as a
    a. Reefer
    b. Key
    c. Lid

—Written and administered by students from Callanan Junior High School, Des Moines, Iowa

Key to S.H.A.F.T. Test

1. c	9. c	16. b
2. b	10. b	17. a
3. b	11. b	18. a
4. c	12. b	19. b
5. a	13. c	20. c
6. b	14. a	21. a
7. c	15. b	22. c
8. c		

How To Score

0- 4 correct—Nothing
5-10 correct—Yellow armband
11-16 correct—Red armband
17-22 correct—Black armband

Wrap-Up

Speaker: Dr. Dwight Allen, dean, School of Education, University of Massachusetts, Amherst, Massachusetts.

Testing is a part of the mindless process of education that really is suited to the simpler society of the past. Education is out of date sociologically, psychologically, and physiologically, but I would rather work within the system than burn it down. Either way, traumatic change will hurt people—by and large, the wrong people. If you were to close down all the educational systems in this country, the people in the upper middle class would pretty much do it on their own, but the people who most need to gain access to education would be done in.

This doesn't mean that you shouldn't do in a particular test, because some need to be done in. For example, if you are like most audiences, only about 20 percent of you can name the capitals of North and South Dakota and North and South Carolina—except for the fifth grade teachers. The real standard for an educated person is not being able to name states and their capitals but being able to use an atlas to find them when one needs to. When I suggested to a fifth grade teacher that she let her kids use an atlas on the next test, she said, "Oh, no! You couldn't do that. They'd all get it right."

Our whole testing program is oriented toward a normative assumption—an upper and a lower half. I won't buy into any system that requires someone to fall off the bottom, because this isn't my view of what education is about. An educational system is needed that will assure a win/win proposition for children instead of the present zero/win game.

Very few of us in education ever know when we're done, despite all the tests we give. If a student seems to be about finished, we enrich him, and in the name of high standards we go about it all backwards. Kids soon learn that rewards are given not for achievement, but for keeping the seat warm, being good, and working hard. If a kid comes in with a theme, throws it on the teacher's desk, and says, "Here's something I whipped off in 20 minutes. I hope you like it," the teacher says, "You should have worked harder." If the same kid brings in the same paper and says, "Here's the eighth draft. I wish I had time for a ninth," she says, "Nice boy, Johnny. You are working hard."

A revitalization of the testing program has to be thought of in terms of the broader objectives of education. The problem with this, however, is that anytime one wants to avoid having to do something, he says, "Let's stop and get our objectives organized." This is good for at least two years. My objectives are the same as yours. I want all the kids to be constructive members of society, to be self-realized and have lots of skills, and to be happy, healthy, and democratic. The objective of our school system, however, is status. With so many people being anointed first with high school diplomas and now with college degrees, it's getting hard to make that status system stick. Education should reflect a status system that is based on legitimate differences in ability.

Here is an example of the nonnormative kind of education I think we ought to have. Graduate education students at UCLA have to take "Statistics for Teachers." Assume that this course is necessary and that the objective is to make students learn it better. The course is divided into 16 units, each one lasting a week. The first unit is called "Counting." The students start out in large groups, enter supplementary groups, and then join small tutorial sessions, not quitting until they have learned the week's work. They are tested as soon as they learn the material—some after the first hour and others after 15 hours. In the typical college course, students take a midterm at the end of about the fifth week. This is the first chance teachers get to find out what their students are learning. Those for whom the instruction hasn't worked are five weeks behind before anybody gets the first inkling. Tests come at the wrong time, but as long as the system has enough people clinking out at the end, we are not bothered about those who are ground up inside. The "Statistics for Teachers" course produces between 85 and 90 percent A's, proving that anybody can learn almost anything that we teach in school. I am not sure that we are serving anybody, however, by reorganizing a system that's already no good and out of phase with what society needs.

The number one priority for every program in the School of Education at the University of Massachusetts is to combat institutional racism. Our admissions procedure produces a bimodal population of students—the Ivy League-Phi Beta Kappa type in one hump, and in the other, the people who are just smart. We talk to prospective students and admit the ones we like. When we get through with them, no one can tell the two groups apart anymore. This is very necessary if you are eliminating a testing program, for you should never patronize the folks at the bottom by giving them a second-class education. We also think that writing dissertations does not indicate leadership in education and that not all doctoral candidates—whether Ivy League or non-Ivy League types—should have to write dissertations. My point is to find a *different* standard and apply it to *all* students.

One of the problems with testing is that it permits decisions only about individuals and never about institutions. I would like to use testing and evaluation to help me make decisions about our program, but I need information about the experiences that contribute to educational competence and leadership. We offer 16 programs, some self-contained with no core requirements, some that are totally touchy-feely, and others that are dice-them-up-and-competency-base-them. Each can succeed on its own terms and is allowed to do so. We must find out what is good for whom and for what. Aptitude treatment interaction may help us. We also have to begin to look at subcultural differences in a different way to find a unity within a diversity, to appreciate being different, to make our educational system reflect this, and to find standards and ways of checking up that are free of the insidious by-products that the testing program has given us. I want to test the people who can learn from tests and prohibit testing the people for whom learning is obstructed by tests.

This society will never succeed until we recognize that we are part of a multiracial, multiclass, multisex world. We must develop strategies to make people produce on their own good intentions. There must be a renovation of testing and curriculum and an end to the idea that teachers are neutral. So long as teachers have to pretend to be objective, so long as schools can teach only those things that are safe for everybody, education will be unreal. We are living in a complex world, and our school system has to become complex.

This is a condensation of Dr. Allen's speech. The full text is available from the NEA Center for Human Relations; single copies free while they last.

Recommendations

More than 50 small groups worked throughout the conference to discuss the issues and to develop recommendations. The limitations of time did not permit careful consideration of these recommendations by the entire conference. The following list represents a summary of the primary recommendations and concerns of the participants.

To Teachers and Counselors

Forum C recommended that teachers and counselors participate in in-service workshops to increase their awareness of the implications of tests and to learn more about creative ways to help pupils. Forum participants also urged teachers to supplement their pupils' test scores with notes, letters, and anecdotal records.

To State Departments of Education

Forum A recommended that state departments of education stop using the National Teacher Examination to determine certification and employment. *Forum B* recommended that state departments discontinue use of the term *IQ*.

To the Testing Industry

Forum A censured the testing industry for failing to respond to the needs of culturally different children. Forum members urged the industry's test planners to use behavioral objectives developed by teachers, students, and community people; they also recommended that the Educational Testing Service state that the NTE should not be used to determine certification and employment.

To Colleges and Universities

The recommendations of *Forum A* stressed the need for colleges and universities to make teacher training programs more person-oriented and to improve instruction in child development, cultural awareness, and tests and measurements, the latter to deal with the limitations of tests for culturally different learners. Participants called for an end to the use of the Miller Analogies Test, Graduate Record Examination, and National Teacher Examination and urged that colleges and universities not use standardized tests to determine admission, placement, or employment. They called for an end to the term *IQ* and an abolition of group tests of mental ability.

To School Boards and Administrators

Forum A participants called for an immediate end to the use of standardized tests and the removal of standardized and IQ-test scores from cumulative records. They also urged school administrators to involve students, teachers, and community people in decisions about the use of tests and to provide released time and resources for in-service training programs about testing.

Members of *Forum B* urged administrators to stop using the term *IQ* and to abolish all group tests of mental ability.

Participants in *Forum C* asked adminis-

trators to administer standardized tests only if the professional staff and board members understand their limitations and uses and to require in-service training in this area. They called on school boards to formulate policy concerning the right of parents to have access to the complete test records of their children with the assistance of trained personnel. They urged administrators to develop guidelines that restrict the use of group tests to curriculum development, program concerns, and determination of needs for special services; that prohibit the use of test scores to label individuals and institutions; and that require that scores obtained from group tests given to evaluate an institution be removed from individual records and be available for professional use only.

To NEA

Forum A recommended that NEA bring suit, file an *amicus curiae* brief, or provide legal assistance in court suits involving standardized tests and oppose the inclusion of standardized test scores in data banks. Participants asked NEA to recommend that student teachers not be hired if they come from institutions with inadequate instruction in child development, cultural awareness, and tests and measurements. Participants urged that the Association's Resolutions Committee prepare a resolution calling for the discontinuation of standardized tests that are misused and reactivate previously passed resolutions relating to the abuses of tests. Locals were urged to form representative task forces to review the use of tests on the culturally different learners in their communities.

Forum B urged an end to the use of the term *IQ* and abolition of group tests of mental ability.

Forum C recommended that NEA (a) underwrite research with HEW and other organizations to develop tests that are relevant, functional, and helpful to the persons being tested and involve students, school staff, minority professionals, and nonwhite women in all stages of the research; (b) join with state legislative commissions to eliminate use of the NTE and urge the Educational Testing Service to discontinue this test; and (c) join with the American Psychological Association, the American Personnel and Guidance Association, and the American Association of Psychologists for La Raza to demand that test publishers publish instruments that are fair to the testees and to declare a moratorium until this has been achieved.

Forum D recommended that the NEA president and executive secretary and the director of the NEA Center for Human Relations work with other organizations to declare an immediate moratorium on standardized testing, to remain in force until criterion referenced and performance objective tests and valid instruments for testing bilingual children are developed, used, and monitored by Chicano, Black, or First American psychologists. Other groups to be involved include state departments of education, colleges and universities, accreditation associations, organizations of school social workers and psychologists, school boards, and the testing industry.

Forum D members also recommended that NEA declare as policy that an individual's performance on standardized group tests should not be described as IQ and that institutions should discontinue tracking and ability grouping that cause cultural segregation and negative stereotyping of children. These concerns should be communicated to other organizations.

Recommendations of the Chicano Caucus

Be it Resolved, that the NEA, with the approbation of the Bay Area Bilingual Education League (BABEL), print and disseminate the BABEL test and assessment report; that the report be disseminated to all NEA departments, national affiliates, and local and state associations with a directive, sanctioned by the NEA Executive Committee, to appoint

special task groups under the guidance of and according to the specifications recommended by the NEA Chicano Caucus and the NEA First American and Hispanic Task Force, to develop recommendations for teacher training and curriculum development models based on these recommendations. Further, that these recommendations be presented for adoption at the 1972 NEA convention.

Be it Resolved, that the affirmative action program for First American and Hispanic education accepted by the NEA Executive Committee comply with Title VII of the Civil Rights Act of 1964 and that the NEA employ 20 Chicanos in professional positions at the national and regional levels by June 25, 1972.

Be it Further Resolved, that the executive secretary of the NEA report to the 1972 Representative Assembly the progress of the affirmative action program and the implementation of new business items regarding minority employment.

Be it Known, that the Chicano Caucus endorses the principles of the Bay Area Bilingual Education League.

Be it Resolved, that information about the NEA-CHR Chicano Caucus be written in NEA's official publications; that the Chicano and other caucuses be represented at all NEA functions and that the results be publicized; and that the concerns of the Chicano Caucus be publicized after the approval of the chairman of the Chicano Caucus prior to publication.

Final reports of the forum groups raised the question of a total conference action. Although it was not possible to consider the scope of the recommendations, the following recommendation was made and approved by the total conference delegation.

We recommend that a moratorium on standardized tests, as recommended at this conference, be instituted immediately.

The following statement on the testing of bilingual children was submitted to forum groups by representatives of the Bay Area Bilingual Education League. A majority of the forum groups had the opportunity to study the statement and did endorse it.

Testing Bilingual Children

Testing of children whose language is other than standard English with instruments that were developed for users of standard English violates the norm and standardization of these instruments and makes the results questionable. We contend that the use of these instruments with children whose language is other than standard English is invalid.

Sufficient evidence now exists to direct us to the development of criterion-referenced assessment systems as a means of improving the accountability of educational programs. These evaluation processes must correspond to local performance objectives.

The development of valid test instruments for bilingual and bicultural children must be directed by qualified bilingual and bicultural personnel in the educational field or in similar fields, to assure that the test instruments will reflect the values and skills of the ethnic and cultural groups being tested.

Whereas currently used standardized tests measure the potential and ability of neither bilingual nor bicultural children and yet are so used and relied upon to count, place, and track these children, we resolve that such use of standardized tests be immediately discontinued.

—Submitted by the Bay Area Bilingual Education League (BABEL), 1414 Walnut Street, Berkeley, California 94709

APPENDIX

Testing Opinionnaire

A 22-item opinionnaire on testing issues was compiled by the CHR staff and sent to registered participants before the conference to involve them in considering the issues and to obtain some indication of their opinions. A total of 301 participants completed and returned the opinionnaire. Their responses, which were considered neither right nor wrong, were used as a starting point for discussion during the conference.

The average (mean) ratings of participant responses were computed (see Table I). These responses made it clear that participants did not accept tests uncritically and were especially concerned about the ways that testing practices affect minority group children.

Thirty-six percent of the participants indicated that more than 60 percent of all school children are stigmatized by standardized tests; only 5 percent of the participants believed that less than 10 percent of school children are stigmatized by testing.

The last portion of the opinionnaire asked participants to select one of four alternatives for action—eliminate use of tests, intensify efforts to develop culture-free tests, curtail the use of standardized tests except for research purposes, or conduct an intensive educational program to prevent misuse of tests. Nearly half of the participants selected the fourth option. A total of 17 percent recommended eliminating standardized tests or restricting their use to research purposes only. This preconference option is particularly interesting in view of the conference recommendation calling for a moratorium on testing. Although no attempt was made to document the impact of the conference on participants, it could be speculated that the primary outcome was to intensify commitment to stronger action alternatives. This point needs to be tested.

Table I
Mean Ratings – Testing Opinionnaire

5 – Strongly Agree
4 – Slightly Agree
3 – No Opinion or Don't Know
2 – Slightly Disagree
1 – Strongly Disagree

Opinion	*Average Rating* N = 301	*Majority Agreement with Item*
A. Most colleges have developed procedures which minimize discrimination against minority group members on the basis of test scores.	2.14	No
B. Evaluation techniques should be used to provide information or feedback to students and teachers in the development of individualized instruction programs.	4.88	Yes
C. For the most part, tests are taken for granted as measuring instruments and accepted uncritically as the best available instrument at hand.	4.36	Yes
D. Standardized tests are accurate predictors of vocational performance.	1.45	No
E. Standardized tests compensate for language and cultural differences of minority groups.	1.28	No
F. Parents should have access to information regarding their children's scores on standardized tests.	4.16	Yes
G. Primary use for testing should be the development of individualized instruction plans for all students.	4.49	Yes
H. Differences in pupil motivation are adequately compensated for in standardized tests.	1.66	No
I. Standardized tests have successfully bypassed problems of sex role stereotypes.	1.94	No

Opinion	Average Rating N = 301	Majority Agreement with Item
J. IQ tests are not perfectly accurate nor are they a perfect indicator of potential.....	4.74....	Yes
K. IQ tests have been used to deny minority children educational opportunities by labeling them as uneducable, placing them in ii appropriate ability groupings, and special education classes.....	4.56....	Yes
L. The National Teachers Examination (NTE) is an accurate predictor of success in teaching.....	1.71....	No
M. The emphasis for children with limited backgrounds should be placed on current achievement, rather than on intelligence or on assignment to learning groups.....	4.29....	Yes
N. The IQ test is a measure of experience and learning rather than a measure of inborn ability.....	4.31....	Yes
O. Ability grouping is widely approved by school teachers and administrators.....	3.69....	Yes
P. The meaning of test scores is widely misunderstood by teachers, administrators, and the general public.....	4.41....	Yes
Q. Standardized tests are a necessary part of school evaluation.....	2.92....	No
R. Most standardized tests are tests of developed abilities rather than measures of potential.....	4.37....	Yes
S. A major disadvantage for many minority group children is that teaching and testing are usually entirely in "standard" American, a "second language" for many.....	4.57....	Yes
T. A primary outcome of ability grouping is to reinforce positive self-concepts of high scores, but also to reinforce negative self-concepts of low scores.....	4.07....	Yes

Question	Percentage (%)
U. The number of school children stigmatized by standardized tests is probably –	
a. Less than 10%	5%
b. 11–20%	9%
c. 21–40%	29%
d. 41–60%	21%
e. 61–80%	19%
f. 81% or more	17%
	100%
V. Given the possible negative effects of standardized tests, which of the following actions do you believe should be taken?	
a. Eliminate the use of standardized tests entirely	6%
b. Intensify efforts to develop culture-free tests	34%
c. Curtail the use of standardized tests except for research purposes	11%
d. Conduct an intensive educational program to prevent misuse of tests	49%
	100%

"It is our feeling that standardized tests are used, actually and intentionally, to discriminate against minority people."

"We feel that a three- to five-year moratorium on group and individual IQ testing should be instituted immediately."

"The use of standardized test scores creates pressure from college through elementary school, and many students are penalized because of their scores."

"Properly used standardized tests can provide a better instructional program by identifying students' strengths and weaknesses."

"We regard children as test scores rather than as human beings."

"Don't vote for any candidate who has no position on bilingual education."

"My daughter doesn't need another slot; being Black is enough."

"We educators keep one-fourth of the America people for 12 years. We could remake this country if we wanted to."

"Evaluators seem to favor the test maker over the test taker."

"If a significant portion of the white society had felt tests to be unfair, the tests would have been changed by now."

"Teacher to student: 'You don't have to take this test, but if you don't, it will go on your record and follow you for life.' "

"All testing, valid or otherwise, should be deemphasized to rid this society of its unwarranted dependence on testing."